NETWORK UNIX 1

RFC 681 NIC 32157 2

INTRODUCTION 3

THE UNIX TIME-SHARING SYSTEM [1] PRESENTS SEVERAL INTERESTING CAPABILITIES AS AN ARPA NETWORK MINI-HOST. IT OFFERS POWERFUL LOCAL PROCESSING FACILITIES IN TERMS OF USER PROGRAMS, SEVERAL COMPILERS, AN EDITOR BASED ON QED, A VERSATILE DOCUMENT PREPARATION SYSTEM, AND AN EFFICIENT FILE SYSTEM FEATURING SOPHISTICATED ACCESS CONTROL, MOUNTABLE AND DE-MOUNTABLE VOLUMES, AND A UNIFIED TREATMENT OF PERIPHERALS AS SPECIAL FILES. 3a

THE NETWORK CONTROL PROGRAM (NCP), IS INTEGRATED WITHIN THE UNIX FILE SYSTEM. NETWORK CONNECTIONS ARE TREATED AS SPECIAL FILES WHICH CAN BE ACCESSED THROUGH STANDARD UNIX I/O CALLS; VIZ. READ, WRITE, OPEN, CLOSE. SPECIAL FILES HAVE DIRECTORY ENTRIES SIMILAR TO NORMAL FILES EXCEPT THAT CERTAIN FLAG BITS ARE SET. THESE FLAG BITS CAUSE SYSTEM I/O ROUTINES TO TAKE SPECIAL ACTION. IN UNIX, SPECIAL FILES SIGNIFY PERIPHERAL DEVICES. FOR EXAMPLE, I/O TRANSACTION WITH MAGTAPE ZERO WOULD BE ACCOMPLISHED BY ACCESSING THE SPECIAL FILE, "/DEV/MT0". FOR THE UNIX NETWORK SYSTEM, ADDITIONAL SPECIAL FILES WERE CREATED EACH OF WHICH SPECIFIES A HOST ON THE ARPA NETWORK. FOR EXAMPLE "/DEV/NET/HARV" REPRESENTS THE PDP-10 AT HARVARD. THIS SIMPLE ACCESS MECHANISM, THROUGH THE FILING SYSTEM, ALLOWS STANDARD ARPA PROTOCOLS SUCH AS TELNET AND FTP TO BE IMPLEMENTED AS SWAPPABLE USER PROGRAMS, RESIDENT ONLY WHEN NEEDED. FURTHERMORE, A USER MAY WRITE HIS OWN PROGRAMS TO COMMUNICATE WITH THESE SPECIAL FILES JUST AS THE TELNET PROGRAM DOES. THE SAMPLE PROGRAM FOUND BELOW DEPICTS THE ESSENTIALS OF NETWORKING FROM UNIX. 3b

STANDARD I/O 4

TO PRESENT THE BASIC PROPERTIES OF UNIX I/O, THE READ, WRITE, OPEN, AND CLOSE FUNCTION CALLS ARE SUMMARIZED BELOW. EACH CALL MAY RESULT IN AN ERROR CODE OF MINUS ONE. 4a

TO MANIPULATE AN EXISTING FILE WITH READS OR WRITES, IT MUST FIRST BE OPENED WITH THE FOLLOWING CALL: 4b

```
FILEDES = OPEN( "ANYFILENAME",FLAG )
```
4b1

WHERE "ANYFILENAME" IS THE ARBITRARY NAME OF THE FILE TO BE OPENED. THE SECOND PARAMETER INDICATES WHETHER THE FILE IS TO BE READ, WRITTEN, OR UPDATED. THE RETURNED VALUE "FILEDES", IS CALLED A FILE DESCRIPTOR. IT IS AN INTEGER USED TO IDENTIFY THE FILE IN SUBSEQUENT CALLS TO READ AND WRITE. 4c

ONCE A FILE HAS BEEN OPENED, THE FOLLOWING CALLS MAY BE USED: 4d

```
NBYTES = READ( FILEDES,BUFFER,COUNT );
NBYTES = WRITE( FILEDES,BUFFER,COUNT );
```
4d1

COUNT IS THE NUMBER OF BYTES TO BE TRANSMITTED BETWEEN THE FILE REPRESENTED BY 'FILEDES' AND THE BYTE ARRAY REPRESENTED BY 'BUFFER'. NBYTES IS THE NUMBER ACTUALLY TRANSMITTED. FOR THE READ CALL, 'NBYTES' MAY BE ZERO TO INDICATE THE END OF FILE; IN EITHER CASE, MINUS ONE WILL BE RETURNED IF THERE WAS AN ERROR. 4e

FOR EACH OPEN FILE, THE SYSTEM MAINTAINS A POINTER TO THE NEXT BYTE TO BE READ OR WRITTEN. IF N BYTES ARE TRANSMITTED, THE POINTER ADVANCES N BYTES. DATA WRITTEN TO A FILE AFFECT ONLY THOSE BYTES IN THE FILE WHICH ARE INDICATED BY THE POSITION OF THE WRITE POINTER AND THE COUNT; NO OTHER PART OF THE FILE IS CHANGED. IF THE SYSTEM POINTER INDICATES THAT ANY BYTES BEING WRITTEN WOULD LIE BEYOND THE END OF THE FILE, THE FILE IS ENLARGED AS NEEDED. 4f

ONCE THE USER HAS FINISHED PROCESSING A FILE, IT SHOULD BE CLOSED. THIS IS AFFECTED WITH THE FOLLOWING CALL: 4g

```
CLOSE( FILEDES );
```
4g1

ALTHOUGH IT IS NOT ABSOLUTELY NECESSARY TO DO A SPECIFIC CLOSE ON A FILE WHEN FINISHED, (THE SYSTEM CLOSES ALL FILES WHEN A PROGRAM EXITS), IT IS A GOOD PRACTICE, SINCE THE USER IS ALLOWED ONLY SIXTEEN OPEN FILES. 4h

THERE ARE SEVERAL ADDITIONAL SYSTEM CALLS RELATED TO I/O WHICH WILL NOT BE DISCUSSED IN DETAIL. A FEW OF THE MORE NOTABLE ONES ALLOW THE USER TO: GET THE STATUS OF A FILE, CHANGE THE PROTECTION OR OWNERSHIP OF A FILE, CREATE A FILE, CREATE A DIRECTORY, MAKE A LINK TO AN EXISTING FILE, AND DELETE A

FILE. FOR FURTHER INFORMATION CONCERNING THE DIFFERENT I/O CALLS THE READER IS DIRECTED TO THE UNIX PROGRAMMER'S MANUAL, FIFTH EDITION, K. THOMPSON, AND D. M. RITCHIE, JUNE 1974. 4i

THE USER COMMUNICATES WITH THE NETWORK VIA THESE SAME SYSTEM CALLS. FOR EXAMPLE, IF ONE WISHED TO CONNECT TO THE THE PDP-10 AT HARVARD, THE FOLLOWING SEQUENCE OF CALLS MIGHT BE USED. 4j

```
FILEDES = OPEN( "/DEV/NET/HARV",2 );
IF( FILEDES < 0 )
   PRINTF(" HARVARD IS DEAD");
ELSE
   WHILE( (NBYTES=READ(FILEDES,BUF,80)) > 0 )
      WRITE( 0,BUF,NBYTES );
```
4j1

THE OPEN INSTRUCTS THE SYSTEM TO OPEN A TELNET CONNECTION TO HARVARD, IF MINUS ONE IS RETURNED, THE PROGRAM PRINTS A MESSAGE AND EXITS, OTHERWISE THE PROGRAM WILL READ ANY BYTES SENT BY HARVARD AND PRINT THEM OUT ON THE CONTROLLING TELETYPE. THIS WILL GO ON UNTIL HARVARD CLOSES THE CONNECTION (READ WILL RETURN MINUS ONE WHEN THE CONNECTION IS CLOSED). 4k

UNIX TELNET 5

IN ORDER TO COMMUNICATE WITH REMOTE HOSTS ON THE ARPA NETWORK, ONE FIRST LOGS IN TO UNIX AS A NORMAL USER. THE USER THEN RUNS A PROGRAM, TELNET, WHICH AFTER ANNOUNCING ITSELF LEAVES HIM WITH SEVERAL OPTIONS. 5a

HE MAY CONTINUE WITH HIS NORMAL UNIX ACTIVITIES. WHEN TELNET SEES A UNIX COMMAND, IT WILL INITIATE THE REQUEST AS A PARALLEL TASK, IN THE SAME MANNER AS THE UNIX COMMAND PROCESSOR (THE SHELL). SINCE THIS MAY BE DONE REGARDLESS OF WHETHER OR NOT A NETWORK CONNECTION IS OPEN, THE USER MAY SIMULTANEOUSLY RECEIVE OUTPUT FROM A FOREIGN HOST'S SERVER TELNET AND CONVERSE WITH THE LOCAL UNIX SYSTEM. 5b

WHEN THE TELNET-USER OPENS A CONNECTION, TELNET ACCEPTS THE HOST NAME AND ANY SPECIAL PARAMETERS, AND DOES AN OPEN ON THE SPECIAL FILE CORRESPONDING TO THAT HOST. WHEN CONTROL IS RETURNED, THE CONNECTION IS OPEN. ANY FURTHER DATA RECEIVED FROM THE TERMINAL NOT CONTAINING ESCAPE CHARACTER IS SENT TO THE

NETWORK FILE. ANY DATA RECEIVED IN RESPONSE TO A READ ON THE NETWORK FILE, IS WRITTEN ON THE USER'S TYPEWRITER. 5c

COMMUNICATION CONTINUES WITH THE HOST UNTIL THE USER WISHES TO CLOSE THE CONNECTION. THE USER SIMPLY MAKES THIS KNOWN TO TELNET VIA A COMMAND, AND TELNET DOES A STANDARD CLOSE ON THE NETWORK FILE. THE NEGOTIATION OF CLOSING THE NETWORK CONNECTION IS LEFT TO THE SYSTEM, FREEING THE USER FOR OTHER COMPUTATIONAL WORK. 5d

THERE IS SOME CHARACTER TRANSLATION AND INVISIBLE CONTROL INFORMATION PASSED BACK AND FORTH BETWEEN THE FOREIGN HOST AND THE TELNET PROCESS. THIS INVOLVES RECOGNITION OF TELNET IACS AND THE TRANSLATION OF CARRIAGE RETURN(CR) AND LINE FEED(LF) TO LINE FEED ON ALL DATA RECEIVED FROM THE NETWORK, AND THE INVERSE TRANSLATION OF LF TO CR LF ON ALL DATA SENT TO THE NETWORK. 5e

NCP STRUCTURE 6

DUE TO THE STRUCTURE OF BOTH THE IMP TO HOST[2] AND HOST TO HOST[3] NETWORK PROTOCOLS, DATA COMES FROM THE NETWORK DESTINED NOT ONLY FOR ONE OF MANY ACTIVE PROCESSES, BUT FOR THE INFORMATION OF THE LOCAL HOST AS A WHOLE. FOR EXAMPLE, NETWORK TRAFFIC SUCH AS A HOST TO HOST RESET, WHICH GENERALLY SIGNALS THAT A FOREIGN HOST HAS COME "ALIVE" MUST BE ACKNOWLEDGED TO LET THAT HOST KNOW THAT THE LOCAL HOST ITSELF IS "ALIVE". THEREFORE, THE LOCAL HOST MUST MONITOR DATA COMING FROM THE NET TO PERFORM NOT ONLY A MESSAGE SWITCHING FUNCTION, WHICH IS THE BULK OF NETWORK TRAFFIC, BUT TO PROVIDE A CONTROL AND STATUS FUNCTION. 6a

FURTHER, WHEN A PERSON ASSOCIATED WITH THE LOCAL HOST WISHES TO CARRY ON A CONVERSATION WITH A NETWORK SERVER, THE INITIAL CONNECTION PROTOCOL[4] MUST BE USED TO PROVIDE A LOGICAL PORT AT EACH SITE FOR SUCCEEDING INFORMATION FLOW. 6b

EXPERIENCE WITH THE ANTS MARK I[5] AND ANTS MARK II[6] SYSTEMS HAS SHOWN THAT THE ABOVE CLASSES OF NETWORK EVENTS ARE RELATIVELY INFREQUENT, AND THAT MOST NETWORK TRAFFIC IS IN TERMS OF USER DATA FLOW AND THE ASSOCIATED FLOW CONTROL(HOST TO HOST ALLOCATES AND IMP TO HOST RFNMS). IT IS ALSO THE CASE THAT THE SOFTWARE REQUIRED TO IMPLEMENT THE STATUS AND CONTROL FUNCTION IS THE BULKIEST PART OF AN NCP. 6c

IN UNIX, THE KERNEL OF THE OPERATING SYSTEM IS CORERESIDENT AND NON-SWAPPABLE. A LARGE KERNEL REDUCES THE MEMORY AVAILABLE FOR USER PROGRAMS. THUS IT IS DESIRABLE TO MINIMIZE THE AMOUNT OF CODE ADDED TO THE BASIC UNIX KERNEL FOR THE NCP. FOR THIS REASON, THE NCP IS IMPLEMENTED IN TWO PARTS. ONE PART IS ROOTED IN THE KERNEL AND MAKES UP THE NON-SWAPPABLE SECTION, ABOUT 3.5K WORDS OF CORE. THE OTHER SECTION (CALLED THE NCP DAEMON) DEALS WITH USER REQUESTS TO OPEN AND CLOSE CONNECTIONS AND HANDLES THE STATUS TRAFFIC DESCRIBED ABOVE. THE NCP DAEMON RUNS AS A SWAPPABLE USER PROCESS OF ABOUT 8.5K WORDS IN SIZE, AND COMMUNICATES WITH THE KERNEL VIA A SPECIAL FILE. 6d

HARDWARE AND SOFTWARE REQUIREMENTS 7

THE NETWORK SOFTWARE FOR UNIX WAS DEVELOPED ON A PDP-11/50, WITH MEMORY MANAGEMENT, TWO RK05 DISK PACKS, TWO NINE TRACK MAGTAPE DRIVES, FOUR DECTAPE DRIVES, 32K WORDS OF CORE, AND THREE TERMINALS. PRESENTLY THIS HAS BEEN EXPANDED TO ENCOMPASS A DH11 TERMINAL MULTIPLEXOR, AN RP03 MOVING HEAD DISK, A TWIN PLATTER RF11 FIXED HEAD DISK, FLOATING POINT, AND 48K OF CORE. USER FILES ARE STORED ON THE RP03. THE RF11 IS USED AS A SWAP DISK AND FOR TEMPORARY FILE STORAGE; ONE RK05 PLATTER CONTAINS THE SYSTEM FILES, AND THE SECOND CONTAINS LOGIN AND ACCOUNTING INFORMATION. IN THE NEAR FUTURE, THE SYSTEM WILL BE EXPANDED TO 128K WORDS OF CORE MEMORY WITH 10 DIAL IN AND 10 HARD WIRED TERMINAL LINES. 7a

THE BASE OPERATING SYSTEM OCCUPIES 24.5K WORDS OF MEMORY. THIS SYSTEM INCLUDES A LARGE NUMBER OF DEVICE DRIVERS, AND ENJOYS A GENEROUS AMOUNT OF SPACE FOR I/O BUFFERS AND SYSTEM TABLES. A MINIMAL SYSTEM WOULD REQUIRE 40K WORDS OF HARDWARE MEMORY. IT SHOULD BE NOTED THAT UNIX ALSO REQUIRES THE MEMORY MANAGEMENT OPTION OFFERED BY DEC TO RUN AT ALL. 7b

THE BASE OPERATING SYSTEM WAS DEVELOPED BY BELL LABORATORIES IN MURRAY HILL, NEW JERSEY. THE BELL INSTALLATION SUPPORTS A HIGH SPEED PAPER TAPE READER-PUNCH, NINE-TRACK MAGNETIC TAPE, AND DECTAPE. BESIDES THE CONSOLE TERMINAL, THERE ARE 14 VARIABLE SPEED COMMUNICATION DATASETS, AND A 201 SERIES DATASET FOR SPOOLING PRINTOUT TO A COMMUNAL LINE PRINTER. THERE ARE ALSO SEVERAL ONE-OF-A-KIND DEVICES INCLUDING A VOICE RESPONSE UNIT, A VOICE SYNTHESIZER, A PHOTOTYPESETTER, A DIGITAL SWITCHING

NETWORK, AND A SATELLITE PDP-11/20 WHICH GENERATES VECTORS, CURVES, AND CHARACTERS FOR A TEKTRONIX 611 STORAGE-TUBE DISPLAY. 7c

RELIABILITY 8

AS OF THIS WRITING, NETWORK UNIX HAS BEEN RUNNING ON A FULL TIME BASIS FOR ABOUT FOUR WEEKS. DURING THAT PERIOD, THERE WERE BETWEEN THREE AND FOUR CRASHES A DAY. THIS IS NOT A VALID INDICATOR BECAUSE MANY OF THE FAILURES WERE DUE TO HARDWARE COMPLICATIONS. MORE RECENTLY THE HARDWARE HAS BEEN RE-CONFIGURED TO IMPROVE RELIABILITY AND THE CRASH RATE HAS BEEN REDUCED TO ONE A DAY WITH A DOWN TIME OF 2-3 MINS. THIS IS EXPECTED TO CONTINUE, BUT THE SAMPLING PERIOD HASNT BEEN LONG ENOUGH FOR ANY DEPENDABLE ANALYSIS. 8a

AVAILABILITY 9

ALTHOUGH THE UNIX NETWORK SOFTWARE WAS DEVELOPED WITHOUT ARPA SUPPORT, THE CENTER FOR ADVANCED COMPUTATION IS WILLING TO PROVIDE IT GRATIS TO THE PEOPLE OF THE ARPA COMMUNITY. 9a

HOWEVER BELL LABORATORIES MUST BE CONTACTED FOR A LISCENSE TO THE BASE SYSTEM ITSELF. BELL'S POLICY IN THE PAST HAS BEEN TO LISCENSE THE SYSTEM TO UNIVERSITIES FOR A NOMINAL FEE, $150.00, AND UNFORTUNATELY FOR A COST OF $20,000.00 TO "NONUNIVERSITY" INSTITUTIONS. 9b

IN THIS LIGHT BELL WAS APPROACHED TO SEE WHAT THEIR REACTION WOULD BE TO AN ARPA NETWORK WIDE LISCENSE, THEY SAID THEY WERE OPEN TO SUGGESTIONS IN THAT AREA. SO SHOULD ENOUGH PEOPLE BECOME INTERESTED, PERHAPS A LESS EXPENSIVE FEE CAN BE NEGOTIATED. 9c

INTERESTED USERS WHO HAVE EITHER SOURCE LISTINGS OR SOURCE FILES INCLUDE: 9d

THE RAND CORPORATION WHICH IS USING OUR IMPLEMENTATION AS A BASIS FOR THEIR OWN VERSION. 9e

LINCOLN LABORATORIES WHICH HAS A SOURCE LISTING TO BE USED AS AN AID IN EVALUATION OF THE UNIX SYSTEM. 9f

THE INCO CORPORATION OF MC LEAN VIR. HAS A LISTING TO HELP IN THE INSTALLATION OF AN NCP INTO DEC'S RSTS OPERATING SYSTEM. 9g

IN ANY CASE WE ARE WILLING TO HELP ANY GROUP WITH ACQUISITION OF A SYSTEM. 9h

FOR FURTHER INFORMATION CONCERNING THE SYSTEM CONTACT: 9h1

STEVE HOLMGREN
210 ADVANCED COMPUTATION BLDG.
UNIVERSITY OF ILLINOIS
URBANA ILLINOIS 61801

(217)-333-8469
OR
HOLMGREN AT BBN

OUTLOOK AND FUTURE PLANS 10

WITH THE ADVENT OF TELNET IN UNIX, CURRENT PLANS ARE TO RUN THE SYSTEM OVER THE NEXT ONE OR TWO MONTHS AND WORK OUT ANY REMAINING BUGS. WHILE THIS IS GOING ON, EXTENSIVE BANDWITH AND LOAD TESTING IS GOING TO TAKE PLACE AND ANY REASONABLE IMPROVEMENTS MADE. 10a

AFTER TELNET HAS PROVED ITSELF RELIABLE, THE OPEN SYSTEM CALL WILL BE EXPANDED TO INCLUDE FURTHER PARAMETERIZATION. THIS PARAMETERIZATION WILL ENCOMPASS CONNECTIONS TO SPECIFIC SOCKETS, SIMPLEX CONNECTIONS BASED ON A SOCKET ALREADY IN USE, AND THE ABILITY TO LISTEN ON A LOCAL SOCKET. 10b

AFTER THOSE EXTENSIONS, NET MAIL, THEN NETWORK FTP AND FINALLY NETWORK RJE WILL BE IMPLEMENTED. ALL WILL RUN AS USER PROGRAMS SO THE KERNEL SYSTEM SIZE WILL NOT INCREASE. 10c

THERE IS ALSO INTEREST IN IMPLEMENTING SOME OF THE PROCEDURE CALL PROTOCOL BEING DEVELOPED BY THE NATIONAL SOFTWARE WORKS, BUT NO DEFINATE PLAN HAVE BEEN MADE. 10d

ACKNOWLEDGEMENTS 11

I AM MUCH INDEBTED TO GARY GROSSMAN WHO PARTICIPATED IN THE DESIGN AND WROTE THE NCP DAEMON; AND TO STEVE BUNCH WHO WAS THE

THE THIRD MEMBER OF OUR DESIGN GROUP AND WROTE THE KERNEL MESSAGE SOFTWARE. 11a

THE THREE OF US ARE PARTICULARLY APPRECIATIVE OF THE CRITICISM AND SUPPORT OF DR. HUGH FOLK, DR. PETER ALSBERG, GREG CHESSON, JOHN MULLEN, KARL KELLEY AND DAVE HEALY. 11b

REFERENCES 12

1. UNIX TIME-SHARING SYSTEM
KEN THOMPSON AND DENNIS RITCHIE
COMMUNICATIONS OF THE ACM
JULY 1974, VOL 17, NUMBER 7 12a

2. SPECIFICATIONS FOR THE INTERCONNECTION OF A
HOST TO AN IMP
REPORT NO. 1822 BOLT BERANEK AND NEWMAN INC.
CHAPTER 3, SYSTEM OPERATION 12b

3. HOST/HOST PROTOCOL FOR THE ARPA NETWORK
ALEX MCKENZIE, BBN
NIC DOCUMENT 8246 12c

4. OFFICIAL INITIAL CONNECIION PROTOCOL
DOCUMENT #2
J. POSTEL, UCLA-NMC
NIC DOCUMENT 7101 12d

5. ANTS MARK I USER'S GUIDE
KARL KELLEY
CENTER FOR ADVANCED COMPUTATION 2/1/74 12e

6. ANTS MARK TWO SYSTEM
KARL KELLEY
CENTER FOR ADVANCED COMPUTATION 1/10/74 12f